JOINT ESTIMATION OF THE ROOM GEOMETRY AND MODES WITH COMPRESSED SENSING

Helena Peić Tukuljac, Thach Pham Vu,
Hervé Lissek and Pierre Vandergheynst

April 19, 2018
JOINT ESTIMATION OF THE ROOM GEOMETRY AND MODES WITH COMPRESSED SENSING

Content Atlas

Spatio-temporal sampling relations [1]  
Room transfer function [2]  
Room modes [3]

Introduction

Contribution

Spatio-temporal sampling relations

Room transfer function

Room modes

Plane waves and wave vectors

Acoustical k-space


helena.peictukuljac@epfl.ch
Spatio-Temporal Sampling Relations

Sampling steps:
\[ \Delta t, \Delta x, \Delta y, \Delta z \]

Choosing \( \Delta t \) is easy: [1]
\[ \omega \geq 2\omega_{\text{cutoff}} \]

Sampling frequencies:
\[ \varphi_x = \frac{2\pi}{\Delta x}, \varphi_y = \frac{2\pi}{\Delta y}, \varphi_z = \frac{2\pi}{\Delta z} \]

Two points of view on spatial sampling:

View #1: Plenacoustic function
\[ \varphi_y = 0 \text{ m}^{-1}, \quad \varphi_y = 7 \text{ m}^{-1} \]
\[ \omega = \frac{2\pi}{\Delta t}, \varphi_x = \frac{2\pi}{\Delta x}, \varphi_y = \frac{2\pi}{\Delta y}, \varphi_z = \frac{2\pi}{\Delta z} \]

View #2: Courant–Friedrichs–Lewy condition [4]
If a wave is moving across a discrete spatial grid and we want to compute its amplitude at discrete time steps of equal duration, then this duration must be less than the time for the wave to travel to adjacent grid points.

Connected plane wave sparsity and compressed sensing in 2014

Can we introduce a parametric solution and some assumptions that will reduce the complexity of the problem?


helena.peictukuljac@epfl.ch
Wave equation and its parametrized solution

\[
\Delta p(t, X) - \frac{1}{c^2} \frac{\partial^2}{\partial^2 t} p(t, X) = 0
\]

\[
p(t, X) = \sum_{q \in \mathbb{I}} A_q \Phi_q(X) g_q(t)
\]

Spatial dependency:
\[
\Phi_q(X) \approx \sum_{r=1}^{R} a_{q,r} e^{i k_{q,r} \cdot X}
\]

Temporal dependency:
\[
g_q(t) = e^{i k_{q} c t} \quad k_{q} = \frac{\omega_q - j \xi_q}{c}
\]

\[
p(t, X) = \sum_{q,r} \alpha_{q,r} e^{i (k_{q} c t + k_{q,r} \cdot X)}
\]

The goal: Fast and efficient parameter learning from a small number of microphone measurements

\[
p(t, X) = \sum_{q,r} \alpha_{q,r} e^{i ((\omega_q - j \xi_q) t + k_{q,r} \cdot X)}
\]
A set of assumptions for a well-posed problem

What natural assumptions can we introduce to reduce the complexity of the mathematical model at a low cost of approximation losses?

**Assumption #1**: simple (rectangular) room geometries $L_x \times L_y \times L_z$

Axial modes imply room shape

Assumption #2: lightly damped rooms

$$\mathbf{k}_q = \begin{bmatrix} k_x & k_y & k_z \end{bmatrix}^T$$

$$\mathbf{k}_q = \frac{\mathbf{w}_q - j\xi_q}{c}$$

$$\xi \ll \omega \quad |\mathbf{k}| \approx \frac{\omega}{c} = \sqrt{k_x^2 + k_y^2 + k_z^2}$$

Separability of spatial and temporal parameter estimation:
1. temporal: estimate and build a ball $r = \frac{\omega}{c}$
2. spatial: estimate the direction of $\mathbf{k}$

**Assumption #3**: low frequency domain analysis

Spherical Search space

Room mode sparsity

## Three points of view on room modes

<table>
<thead>
<tr>
<th></th>
<th>Plane waves</th>
<th>Pressure isosurfaces</th>
<th>Wave vectors in search space $r = \frac{\omega}{c}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>x-axial mode</strong></td>
<td><img src="image1" alt="x-axial mode" /></td>
<td><img src="image2" alt="Pressure isosurface" /></td>
<td><img src="image3" alt="Wave vector" /></td>
</tr>
<tr>
<td><strong>xy-tangential mode</strong></td>
<td><img src="image4" alt="xy-tangential mode" /></td>
<td><img src="image5" alt="Pressure isosurface" /></td>
<td><img src="image6" alt="Wave vector" /></td>
</tr>
<tr>
<td><strong>oblique mode</strong></td>
<td><img src="image7" alt="oblique mode" /></td>
<td><img src="image8" alt="Pressure isosurface" /></td>
<td><img src="image9" alt="Wave vector" /></td>
</tr>
</tbody>
</table>

helena.peictukuljac@epfl.ch
Focusing on the low frequencies and room mode assignment

**Sound pressure isosurfaces:**
- Axial modes
- Tangential modes
- Oblique modes

Room Transfer Function - Individual Components


helena.peictukuljac@epfl.ch
ReSEMbLE algorithm

How to choose the partitioning frequency $f_p (\omega_p)$?

Based on the approximate room size:
Such that we capture all the basic axial modes

$$\omega_c = \frac{\pi c}{L_{\text{min}}}$$

$\omega_{[1,0,0]} = \frac{\pi c}{L_x}$

$\omega_{[0,1,0]} = \frac{\pi c}{L_y}$

$\omega_{[0,0,1]} = \frac{\pi c}{L_z}$

helena.peictukuljac@epfl.ch
ReSEMbIE algorithm

Temporal dictionary: \[ \theta[i] = e^{j\omega_n[i]t}e^{j\omega_n[i]t} \]
Spatio-temporal dictionary: \[ \Sigma[:, i] = e^{j\omega_n[i]t}e^{j\omega_n[i]t}X_k[i] \]

procedure ReSEMbIE(R, X)
Separate the measurements with \( f_{ij} \): \( R = R^l + R^h \).

for \( i_t \in \{1, ..., N_t\} \) do
  step 1: estimate \( (\alpha_{i_t}, \xi_{i_t}) \) from \( R_{i_t} \)
  step 2: estimate \( k_{i_t} \) from \( R_{i_t}^h \)
  step 3: compute new residual \( R_{i_t+1}^l \)
end for

Recover the room size \( L_x, L_y, L_z \) from basic axial room modes and form the regular wave vector grid.

for \( i_h \in \{N_t + 1, ..., N\} \) do
  step 1: get \( \omega_{i_t} \) and \( k_{i_t} \) from the wave vector grid
  step 2: estimate \( \xi_{i_t} \) from \( R_{i_t}^h \)
  step 3: compute new residual \( R_{i_t+1}^h \)
end for

Estimate the expansion coefficients \( \{\alpha\}^{N}_{i=1, t=1} \) using least square approach.
end procedure

helena.peictukuljac@epfl.ch

Structured group sparsity

Separability of spatial and temporal parameter estimation:

1. temporal: estimate $\omega$ and build a ball $r = \frac{\omega}{c}$

2. spatial: estimate the direction of $k$
k-space estimation results

Why the position of points deviates the most over z-axis? This is why the x-, y- and xy-modes are precise and z-, xz-, xy- and xyz-modes are far off from the ground truth.
ReSEMblE algorithm

Imly the wave vectors from the reconstructed room shape:

\[(n_x, n_y, n_z) \in \mathcal{N}_0^3 \setminus (0, 0, 0)\]

\[k_x = n_x \frac{\pi}{L_x} \quad k_y = n_y \frac{\pi}{L_y} \quad k_z = n_z \frac{\pi}{L_z}\]

After applying the high part of the algorithm, the Pearson correlation coefficient that the approximation is good (e.g. 82% for only 19-microphone setting and \(f_c = 200\)Hz), but it should be further improved once the deviation of the wave vectors is efficiently characterized.
In the spirit of **open research** and **acoustic data augmentation**
Thank you for your time and attention!


helena.peictukuljac@epfl.ch